UNITED STATES PATENT APPLICATION

FOR

DYNAMIC MULTI-HOP INGRESS TO EGRESS L2TP TUNNEL MAPPING

INVENTORS:

Aravind Sitaraman, a citizen of India
Purnam Sheth, a citizen of Canada

ASSIGNED TO:

Cisco Technology, Inc., a California Corporation

PREPARED BY:

THELEN, REID & PRIEST LLP
P.O. BOX 640640
SAN JOSE, CA 95164-0640
TELEPHONE: (408) 292-5800
FAX: (408) 287-8040

Attorney Docket Number: CISCO-3287

Client Docket Number: CISCO-3287

SPECIFICATION

TITLE OF INVENTION

DYNAMIC MULTI-HOP INGRESS TO EGRESS L2TP TUNNEL MAPPING

Cross Reference to Related Applications

This application is related to the following:

U.S. Patent Application Serial No. 09/488,394, filed January 20, 2000 in the name of inventors Aravind Sitaraman, Aziz Abdul, Bernard Janes, Dennis Cox, John Joyce, Peter Heitman, Shujin Zhang and Rene Tio, entitled "System and Method for Identifying a Subscriber for Connection to a Communication Network", commonly assigned herewith.

U.S. Patent Application Serial No. 09/488,395, filed January 20, 2000 in the name of inventors Aravind Sitaraman, Dennis Cox, John Joyce and Shujin Zhang, entitled "System and Method for Determining Subscriber Information", commonly assigned herewith.

U.S. Patent Application Serial No. 09/712,005, filed November 13, 2000 in the name of inventors Purnam Sheth, Aravind Sitaraman, Charles Yager and Gregory Burns, entitled "PPP/L2TP Domain Name Pre-Authorization", commonly assigned herewith.

U.S. Patent Application Serial No. 09/712,780, filed November 13, 2000 in the name of inventors Purnam Sheth, Aravind Sitaraman, Charles Yager and Gregory Burns, entitled "PPP Domain Name and L2TP Tunnel Selection Configuration Override", commonly assigned herewith.

U.S. Patent Application Serial No. _____, filed February 27, 2001 in the name of inventors Aravind Sitaraman and Purnam Sheth, entitled "Load Sharing Between L2TP Tunnels", commonly assigned herewith.

FIELD OF THE INVENTION

The present invention relates to the field of data communications. More particularly, the present invention relates to a system and method for dynamic multi-hop ingress to egress L2TP tunnel mapping.

BACKGROUND OF THE INVENTION

Computer networking capabilities of a home personal computer (PC) are typically provided by telephone companies (Telcos) or commercial Internet Service Providers (ISPs) who operate network access points along the information superhighway. It is through these network access points that the user is able to connect with public domains, such as the Internet, and private domains, such as an intra-company computer network of the user's employer.

In the wholesale Internet access environment, the network access provider (NAP) and the network service provider (NSP) are not necessarily the same entity. Telcos and other wholesale ISPs are typical NAPs, who operate gateways (network access servers, access routers, or the like) in their points of presence (PoPs), and provide local loop access services to PCs. NSPs are typically the customers of NAPs, who are allowed to

use the NAP's gateways to provide their Internet Protocol (IP)-based services, such as Internet access, network access, or voice over IP (VoIP) services to the PCs.

Figure 1 illustrates Layer 2 Tunneling Protocol (L2TP). L2TP tunneling is a common service architecture for Point-to-Point Protocol (PPP) clients currently available at NAPs. In the typical L2TP tunneling environment, a PC 100 of a PPP client 105 starts a PPP session by dialing into a L2TP access concentrator (LAC) 110 located at the NAP's point of presence (PoP). The LAC 110 exchanges PPP messages with the client's PC 100 and communicates with a L2TP network server (LNS) 115 of a remote domain 120 such as an ISP or a private company. The LNS 115 is typically a home gateway (HGW) of the remote domain 120. The communication between the LAC 110 and the LNS 115 is by way of L2TP requests and responses. When a L2TP tunnel 125 is set up, the LAC 110 forwards the PPP session over the L2TP tunnel 125 to the LNS 115. Data packets in the PPP session are encapsulated into L2TP frames that are destined for the IP address of the LNS 115.

The LNS 115 is a termination point of the L2TP tunnel 125. The LNS 115 accepts L2TP frames, strips the L2TP encapsulation, and processes the incoming PPP frames for the appropriate interface. The PPP frames are processed and passed to higher layer protocols, i.e., the PPP session is terminated at the LNS 115. The PPP session termination requires and includes user authentication via a Remote Authentication Dial-In User Service (RADIUS) or other means. An authenticated PPP client then receives an IP address, a Domain Name System (DNS) address, and IP-based services that the client

contracted. These are forwarded back to the client over the L2TP tunnel 125 through the LAC 110.

The L2TP passes protocol-level (or Data Link-level) packets through the virtual tunnel between the endpoints of a point-to-point connection, i.e., the client's PC 100 and the LNS 115. The L2TP is suitable for virtual private networking (VPN), in which users can dial into a NAP's network access server and join a private (typically corporate) network that is remote from the NAP's PoP.

Figure 2 is a block diagram that illustrates a L2TP service architecture in which the NAP and NSP are different entities. Users A 200 and C 205 represent authorized users of a network at remote domain 210 and users B 215 and D 220 represent authorized users of a network at remote domain 225. A router may operate as both a LNS to given set of LACs and simultaneously as a LAC to a given LNS. When configured to operate in this way, the router is called a "multi-hop node" or "egress LAC". In FIG. 2, NAP 230 provides network access to user C 205 and user D 220 and ingress LAC 235 tunnels PPP sessions via ingress tunnels 240 and 270 to egress LACs 245 and 265, respectively. NAP 250 provides network access to user A 200 and user B 215 and ingress LAC 255 tunnels PPP sessions via ingress tunnels 260 and 275 to egress LACs 265 and 245, respectively. Egress LAC 245 aggregates tunnels 240 and 275 into a single egress tunnel 280 to the LNS 285 at remote domain 210. Egress LAC 265 aggregates ingress tunnels 260 and 270 into a single egress tunnel 290 to the LNS 295 at remote domain 225. NSP 203 may also provide services for other remote domains (not shown in FIG. 2).

A typical service level agreement (SLA) specifies a minimum bandwidth to be provided during specified times of the day or week together with pricing information for the provision of such services. As shown in FIG. 2, NAP 250 and NAP 230 have separate SLAs 207, 213 with NSP 203. Additionally, NSP 203 has a third SLA 217 with remote domain 225 and a fourth SLA 223 with remote domain 210. However, sessions utilizing egress tunnel 290 are accorded the same level of service for the interface to LNS 295, and sessions utilizing egress tunnel 280 are accorded the same level of service for the interface to LNS 285.

NSPs typically aggregate a relatively high number of ingress tunnels from NAPs to a relatively low number of tunnels to a remote domain, and egress LACs are statically allocated to particular remote domains based upon expected network usage. As shown in FIG. 2, NSP 203 allocates a separate egress LAC for each remote domain. Egress LAC 265 is allocated for remote domain 295 and egress LAC 245 is allocated for remote domain 210. This static architecture allows for a limited increase in the number of subscribers using a particular egress LAC without reconfiguring the network. For example, if the subscribers using NAP 230 to access remote domain 210 increases beyond the capacity of egress LAC 245 to provide a minimum level of service, the network may have to be reconfigured by, for example, adding another egress LAC at NSP 203.

This approach has a number of drawbacks. First, placing all L2TP sessions for a particular remote domain on a single egress tunnel accords the same level of service to all sessions using that egress tunnel, regardless of any SLAs. Second, tying particular egress LACs to particular remote domains reduces scalability by requiring manual reconfiguration whenever an egress LAC becomes overutilized. Third, static configuration may allow some egress LACs to be underutilized while simultaneously allowing other egress LACs to be overutilized, resulting in a relatively inefficient utilization of resources.

What is needed is a solution that enables differentiated service for sessions using an egress tunnel. A further need exists for such a solution that minimizes the amount of network engineering required by NAPs and NSPs. Yet another need exists for such a solution that allows a service provider to share the same egress LAC across multiple remote domains while maintaining SLAs, providing relatively efficient utilization of resources.

BRIEF DESCRIPTION OF THE INVENTION

A method for dynamic ingress to egress tunnel mapping from a first communication network to a second communication network includes receiving a tunneled communication from a subscriber using the first communication network, determining egress tunnel selection criteria for the tunneled communication, selecting one of at least one egress tunnel based on the egress tunnel selection criteria and forwarding the tunneled communication on the selected egress tunnel. The egress tunnel selection criteria indicate the basis for selecting one of the egress tunnels. An apparatus for dynamic ingress to egress tunnel mapping from a first communication network to a second communication network includes a receiving interface to receive a tunneled communication from a subscriber using the first communication network, an egress tunnel selection criteria determiner to determine egress tunnel selection criteria for the tunneled communication, an egress tunnel selector to select one of at least one egress tunnel based on the egress tunnel selection criteria and a session forwarder to forward the tunneled communication on the selected egress tunnel. In one aspect of the invention, tunnel mapping is performed between Layer 2 Tunneling Protocol (L2TP) ingress and egress tunnels.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated into and constitute a part of this specification, illustrate one or more embodiments of the present invention and, together with the detailed description, serve to explain the principles and implementations of the invention.

In the drawings:

FIG. 1 is a block diagram that illustrates a L2TP tunnel and how a user typically connects to a remote domain.

FIG. 2 is a block diagram that illustrates a L2TP service architecture in which the NAP and NSP are different entities.

FIG. 3 is a block diagram that illustrates a system for dynamic ingress to egress tunnel mapping from a first communication network to a second communication network in accordance with one embodiment of the present invention.

FIG. 4 is a block diagram that illustrates an apparatus for dynamic ingress to egress tunnel mapping from a first communication network to a second communication network in accordance with one embodiment of the present invention.

FIG. 5 is a block diagram that illustrates an apparatus for dynamic ingress to egress tunnel mapping from a first communication network to a second communication network in accordance with one embodiment of the present invention.

FIG. 6 is a flow diagram that illustrates a method for dynamic ingress to egress tunnel mapping from a first communication network to a second communication network in accordance with one embodiment of the present invention.

FIG. 7 is a flow diagram that illustrates a method for selecting an egress tunnel based upon an ingress tunnel ID accordance with one embodiment of the present invention.

FIG. 8 is a flow diagram that illustrates a method for selecting an egress tunnel based upon a subscriber domain in accordance with one embodiment of the present invention.

FIG. 9 is a flow diagram that illustrates a method for selecting an egress tunnel based upon Internet Protocol (IP) packet header Type of Service (ToS) bits in accordance with one embodiment of the present invention.

FIG. 10 is a flow diagram that illustrates a method for selecting an egress tunnel based upon a Virtual Path Identifier (VPI) / Virtual Circuit Identifier (VCI) pair in accordance with one embodiment of the present invention.

FIG. 11 is a flow diagram that illustrates a method for selecting an egress tunnel based upon a pseudo-random process in accordance with one embodiment of the present invention.

FIG. 12 is a flow diagram that illustrates a method for selecting an egress tunnel based upon the time at which a session is received in accordance with one embodiment of the present invention.

FIG. 14 is a flow diagram that illustrates a method for selecting an egress tunnel based upon available bandwidth and an ingress tunnel QoS in accordance with one embodiment of the present invention.

FIG. 15 is a flow diagram that illustrates a method for selecting an egress tunnel based upon available bandwidth and IP header ToS bits in accordance with one embodiment of the present invention.

FIG. 16 is a flow diagram that illustrates a method for selecting an egress tunnel based upon available bandwidth and a VPI/VCI pair in accordance with one embodiment of the present invention.

FIG. 17 is a flow diagram that illustrates a method for selecting an egress tunnel to even the distribution of sessions among egress tunnels in accordance with one embodiment of the present invention.

FIG. 18 is a flow diagram that illustrates a method for diverting tunneled sessions away from an overloaded egress LAC in accordance with one embodiment of the present invention.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

Embodiments of the present invention are described herein in the context of a system and method for dynamic multi-hop ingress to egress L2TP tunnel mapping. Those of ordinary skill in the art will realize that the following detailed description of the present invention is illustrative only and is not intended to be in any way limiting. Other embodiments of the present invention will readily suggest themselves to such skilled persons having the benefit of this disclosure. Reference will now be made in detail to implementations of the present invention as illustrated in the accompanying drawings. The same reference indicators will be used throughout the drawings and the following detailed description to refer to the same or like parts.

In the interest of clarity, not all of the routine features of the implementations described herein are shown and described. It will, of course, be appreciated that in the development of any such actual implementation, numerous implementation-specific decisions must be made in order to achieve the developer's specific goals, such as compliance with application- and business-related constraints, and that these specific goals will vary from one implementation to another and from one developer to another. Moreover, it will be appreciated that such a development effort might be complex and time-consuming, but would nevertheless be a routine undertaking of engineering for those of ordinary skill in the art having the benefit of this disclosure.

In the context of the present invention, the term "network" includes local area networks, wide area networks, the Internet, cable television systems, telephone systems, wireless telecommunications systems, fiber optic networks, ATM networks, frame relay networks, satellite communications systems, and the like. Such networks are well known in the art and consequently are not further described here.

In accordance with one embodiment of the present invention, the components, processes and/or data structures may be implemented using C or C++ programs running on high performance computers (such as an Enterprise 2000™ server running Sun Solaris™ as its operating system. The Enterprise 2000™ server and Sun Solaris™ operating system are products available from Sun Microsystems, Inc. of Mountain View, California). Different implementations may be used and may include other types of operating systems, computing platforms, computer programs, firmware, computer languages and/or general purpose machines. In addition, those of ordinary skill in the art will recognize that devices of a less general purpose nature, such as hardwired devices, field programmable gate arrays (FPGAs), application specific integrated circuits (ASICs), or the like, may also be used without departing from the scope and spirit of the inventive concepts disclosed herein.

The authentication, authorization and accounting (AAA) service performs user authentication, user authorization and user accounting functions. It may be a Cisco ACS™ product such as Cisco Secure™, available from Cisco Systems, Inc. of San Jose, California, or an equivalent product. In accordance with a presently preferred

embodiment of the present invention, the Remote Authentication Dial-In User Service (RADIUS) protocol is used as the communication protocol for carrying AAA information. RADIUS is an Internet standard track protocol for carrying authentication, authorization, accounting and configuration information between devices that desire to authenticate their links and a shared AAA or AAA proxy service. Those of ordinary skill in the art will realize that other authentication protocols such as TACACS+ or DIAMETER can be used as acceptable authentication communications links between the various communications devices that encompass the data communication network and still be within the inventive concepts disclosed herein.

According to embodiments of the present invention, a network tunnel switch such as an Egress LAC or multi-hop node maps sessions in ingress tunnels to a particular egress tunnel that terminates at a remote domain. Egress tunnel selection may be based on a static tunnel-mapping algorithm, in which a session in an ingress tunnel is allocated to an egress tunnel without regard to the current capacity of each egress tunnel to the remote domain. Alternatively, a load factor for available egress tunnels to the remote domain may be considered in determining which egress tunnel to use. The load factor may be based upon parameters such as ingress tunnel ID, subscriber domain, time of day, day of week, quality of service level and subscribed network bandwidth.

Figure 3 is a block diagram that illustrates a system for dynamic ingress to egress tunnel mapping in accordance with one embodiment of the present invention. CorpA 300 and CorpB 305 are remote domains with respect to NSP 310 and NAPs 312, 314 and 316.

NAP 312 provides network access to user A_{CorpA} 318, an employee of CorpA 300. NAP 312 also provides network access to users B_{Corp} 320, I_{CorpB} 322 and J_{CorpB} 324 via ingress LAC 334. NAP 314 provides network access to users C_{CorpA} 326, D_{CorpA} 328, K_{CorpB} 330 and L_{CorpB} 332 via ingress LAC 336. NAP 316 provides network access to users E_{CorpA} 338, F_{CorpA} 340, G_{CorpA} 342 and H_{CorpA} 344 via ingress LAC 346. NAP 316 also provides network access to users M_{CorpB} 348, N_{CorpB} 350, O_{CorpB} 352 and P_{CorpB} 354 via ingress LAC 356. NAP 312 tunnels sessions to NSP 310 via tunnels 358 and 360. NAP 314 tunnels sessions to NSP 310 via tunnels 362 and 364, and NAP 316 tunnels sessions to NSP 310 via tunnels 366 and 368. Egress LAC 370 at NSP 310 serves as a LNS with respect to ingress LACs 334, 336, 346 and 356, while simultaneously serving as a LAC with respect to LNSs 372, 374 and 376. Egress LAC 370 receives tunneled sessions on ingress tunnels 358, 360, 362, 364, 366 and 368 and maps each session to one of egress tunnels 378, 380, 382 and 384. Tunnel database 386 includes tunnel selection criteria for remote domains 300 and 305. Tunnel database 386 also includes parameters for tunnels originating with egress LAC 370. Each NSP may have more than one egress LAC and each remote domain may have more than one LNS.

Turning now to FIG. 4, a block diagram that illustrates an apparatus for dynamic ingress to egress tunnel mapping in accordance with one embodiment of the present invention is presented. Figure 4 provides more detail for egress LAC 370 in FIG. 3. Egress LAC 400 includes a tunnel selection criteria determiner 405, an available egress tunnel ascertainer 410, a tunnel data request generator 415, a tunnel database 420, an egress tunnel selector 425 and a session forwarder 430. The egress LAC 400 may be

implemented using a Cisco 6400 (or an equivalent), available from Cisco Systems, Inc. of San Jose, California.

The tunnel selection criteria determiner 405 determines the criteria to apply in selecting a tunnel received by receiving interface 435. For example, the egress tunnel selection criteria determiner 405 may indicate that egress tunnels should be selected to evenly distribute sessions regardless of an egress tunnel's available bandwidth, or to select an egress tunnel having the most available bandwidth. The available egress tunnel ascertainer 410 determines which egress tunnels to a particular remote domain are available. The tunnel data request generator 415 generates one or more tunnel data requests to obtain the information requested by the available egress tunnel ascertainer 410. The tunnel database 420 stores parameters for each tunnel originating with LAC 400. The tunnel database 420 also stores tunnel selection criteria for each remote domain. Alternatively, tunnel parameters and tunnel selection criteria may be stored in separate databases. The egress tunnel selector 425 selects one of the egress tunnels leading to the remote domain and session forwarder 430 forwards the session received by receiving interface 435 to the remote domain via the egress tunnel selected by egress tunnel selector 425.

In operation, receiving interface 435 receives a session from a first communication network 440 and forwards the session to tunnel selection criteria determiner 405. The tunnel selection criteria determiner 405 determines the criteria to use in selecting a tunnel. According to a preferred embodiment, the selection criteria for

a particular domain is obtained from a database upon receiving a first session destined for the remote domain. If the tunnel selection criteria are independent of tunnel loading, the egress tunnel selector 425 selects the tunnel indicated by the selection criteria determiner 405. If the tunnel selection determiner 405 requires tunnel-loading parameters, available egress tunnel ascertainer 410 ascertains the available egress tunnels to the remote domain. Egress tunnel selector 425 selects one of the tunnels to the remote domain by applying the tunnel selection criteria to the tunnel parameters and session forwarder 430 forwards the session received by receiving interface 435 to the remote domain via the tunnel selected by egress tunnel selector 425.

Turning now to FIG. 5, a block diagram that illustrates an apparatus for dynamic ingress to egress tunnel mapping in accordance with one embodiment of the present invention is presented. Figure 5 is similar to FIG. 4, except tunnel data is maintained in an external database 500. According to one embodiment of the present invention, the tunnel data is stored in an AAA server 505.

Turning now to FIG. 6, a flow diagram that illustrates a method for dynamic ingress to egress tunnel mapping in accordance with one embodiment of the present invention is presented. At 600, a tunnel database is initialized with tunnel selection criteria for each remote domain and with tunnel parameter values for each egress tunnel. For example, tunnel database 386 in FIG. 3 is initialized with tunnel selection criteria for remote domains 300 and 305. The selection criteria for remote domain 300 may indicate sessions should be allocated randomly, while the selection criteria for domain 305 may

indicate sessions should be allocated based upon available bandwidth. In this example, tunnel database 386 also includes tunnel parameters needed for the particular egress tunnel selection criteria. The selection criteria for remote domain 300 require a list of egress tunnels to select from (e.g. egress tunnels 378 and 380). The selection criteria for remote domain 305 requires additional parameters such as the maximum number of sessions for each egress tunnel and the number of sessions currently allocated to each egress tunnel.

Referring again to FIG. 6, at 605, a session from an ingress LAC is received via an ingress tunnel. At 610, the egress tunnel selection criteria for the remote domain are determined. The selection criteria may be obtained from a local database or from an AAA server or the like. Various types of egress tunnel selection criteria are described below in more detail. At 615, a database is checked to ascertain the available egress tunnels to the remote domain and their respective parameters. For example, any tunnel that is "down," is not operational, or is operating over a certain percentage of its capacity could be considered "not available". One or more available tunnels may then be available at 615. At 620, the egress tunnel selection criteria obtained at reference numeral 610 is applied to the tunnel parameters for the available egress tunnels determined at reference numeral 615 to select an egress tunnel. At 625, the session is forwarded on the selected egress tunnel. Processing continues at 605 when another session is received on an ingress tunnel.

As mentioned above, the egress tunnel selection criteria may indicate that egress tunnels should be selected pseudo-randomly, or with a weighted random selection so that those with more available capacity are selected more often, or with any other suitable selection algorithm as will now be apparent to those of ordinary skill in the art. More examples of selection criteria are presented below with reference to FIGS. 7-17.

In accordance with one embodiment of the present invention, an ingress tunnel ID is used to select an egress tunnel for a session. This is illustrated in FIG. 7. At 700, the ID of the ingress tunnel that includes the received session is determined. At 705, an egress tunnel is selected based upon the ingress tunnel ID. For example, a session in an ingress tunnel associated with a relatively high level of service may be allocated to a special egress tunnel for sessions requiring a similar of service. Alternatively, the session may be allocated to any other egress tunnel capable of providing the required level of service.

In accordance with another embodiment of the present invention, a subscriber domain associated with the received session is used to select an egress tunnel for a session. This is illustrated in FIG. 8. At 800, the subscriber domain for the received session is determined. At 805, an egress tunnel is selected based upon the subscriber domain. For example, subscribers of domain xyz.net may be accorded a relatively high level of service. Thus, a session in an ingress tunnel associated with a subscriber from the xyz.net domain may be allocated to a special egress tunnel for sessions requiring a

similar of service. Alternatively, the session may be allocated to any other egress tunnel capable of providing the required level of service.

In accordance with one embodiment of the present invention, the three Precedence bits (the three highest order or most significant bits of the 8-bit Type of Service (ToS)/Differentiated Services Field of the IP packet header) are used to select an egress tunnel for a session. This is illustrated in FIG. 9. At 900, the ToS bits are examined. At 905, an egress tunnel is selected based upon the ToS bits. For example, a session whose ToS bits indicate a relatively high level of service may be allocated to a special egress tunnel for sessions requiring a similar of service. Alternatively, the session may be allocated to any other egress tunnel capable of providing the required level of service.

Those of ordinary skill in the art will realize that the particular bits used are not particularly critical, for example, the CoS (Class of Services) bits of an IEEE 802.1q frame could be used as could the CoS bits in an Inter-Switch Link (ISL) frame. Other bits or fields could also be designated to carry the type of service information. A 3-bit ToS permits up to 8 levels of service. Larger bit fields would be used if desired.

According to another embodiment of the present invention, egress tunnel selection is based upon a VPI/VCI pair associated with the received session. This is illustrated in FIG. 10. As is known to those of ordinary skill in the art, a virtual circuit or channel (VC) is a logical circuit created to ensure reliable communication between two

network devices. A VC is defined by a Virtual Path Identifier (VPI) / Virtual Channel Identifier (VCI) pair. According to this embodiment of the present invention, at 1000, a VPI/VCI pair associated with the received session is examined. At 1005, an egress tunnel is selected based upon the VPI/VCI pair. For example, a session whose VPI/VCI pair indicates a relatively high level of service may be allocated to a special egress tunnel for sessions requiring a similar of service. Alternatively, the session may be allocated to any other egress tunnel capable of providing the required level of service.

According to another embodiment of the present invention, an egress tunnel is selected based upon a pseudo-random process. This is illustrated in FIG. 11. For example, suppose an egress LAC has ten egress tunnels to a remote domain and each egress tunnel is identified by a number from one to ten. A random number generator is used to generate a number from one to ten and the egress tunnel associated with the generated number is selected.

According to another embodiment of the present invention, the time at which a session is received is used to select an egress tunnel. A remote domain may have multiple LNSs in various parts of the world. The workload for each LNS may vary depending upon the time of day, the day of the week, or both. Thus, allocating a session to an egress tunnel associated with a LNS at an "off-hours" site results in a relatively efficient utilization of resources. This embodiment is described below in more detail with reference to FIG. 12.

Turning now to FIG. 12, a flow diagram that illustrates a method for selecting an egress tunnel based upon the time at which a session is received in accordance with one embodiment of the present invention is presented. At 1200, the time at which the session is received is determined. At 1205, the egress tunnel is selected based upon the time the session is received. For example, suppose an egress LAC that receives a session is in country A and that the remote domain has LNSs in country A and country B. Country A is in a time zone that is 12 hours ahead of country B. The egress LAC receives a session at 10AM on a Monday in the egress LAC's time zone, which is 10PM on a Sunday in Country B. In this case, the LNS in Country A will likely have a higher workload than CorpA's LNS in country B. This information is accounted for in an egress tunnel parameter value.

According to embodiments of the present invention, the available bandwidth of egress tunnels and associated interfaces is used either alone or in combination with other parameters to dynamically select an egress tunnel. The selection criteria are applied to the various egress tunnel parameters to create a loading factor. The egress tunnel with the best loading factor is selected. These embodiments are illustrated below with reference to FIGS. 13-16.

According to one embodiment of the present invention, the available bandwidth of all egress tunnels going to a remote domain is used to select an egress tunnel. This is illustrated in FIG. 13. At 1300, the available bandwidth of all egress tunnels going to the remote domain is determined. At 1305, the egress tunnel having the most available

bandwidth is selected. The available bandwidth may be expressed, by way of example, as the maximum number of sessions for an egress tunnel minus the number of active sessions. The maximum number may be hard-coded, or based on the number of sessions the egress tunnel has recently been able to support while maintaining a certain level of service.

The load factor information may be of a number of types, but essentially is an indicator of the available capacity of the particular egress tunnel, weighted by various egress tunnel parameters. In this manner it is now relatively straightforward to program an egress LAC to load balance among the multiple instances of egress tunnels for a remote domain given in the database.

For example, one way to implement the system illustrated in FIG. 3 is to provision $Corp_A$ 300 with a number of egress tunnels for handling sessions and to set the selection criteria such that the egress tunnel with the most available bandwidth is selected. Suppose each egress tunnel is capable of handling up to "X" sessions and currently has two active sessions. In this case, its load factor is X-2. The egress LAC 370 at the NSP's domain can keep track of all sessions passed to these respective egress tunnels so that it can avoid allocating more sessions than any particular egress tunnel can handle. It can also load balance among multiple instances of the egress tunnels so that the load is shared more or less equally. This can be done pseudo randomly or in any other practical manner. Now the load can be balanced and egress tunnels that are unable

to handle sessions can be avoided to minimize the risk of a dropped tunnel, and to minimize the risk of violating a SLA.

According to one embodiment of the present invention, the available bandwidth is weighted by the processing capacity of the LNS CPU. For example, suppose egress tunnel A is currently handling 10 sessions and is associated with a remote domain LNS having a relatively high capacity CPU. Suppose also that egress tunnel B is currently handling the same number of sessions and is associated with a remote domain LNS having a relatively low capacity CPU. In this case, egress tunnel A is better equipped to handle an additional session than egress tunnel B. This information is accounted for in the egress tunnel parameters for egress tunnel A and egress tunnel B.

According to another embodiment of the present invention, an egress tunnel is selected based upon available bandwidth and an ingress tunnel QoS. This is illustrated in FIG. 14. At 1400, the QoS of the ingress tunnel that includes the received session is determined. The QoS of ingress tunnels may be stored, by way of example, in a tunnel database that is indexed by a tunnel ID. At 1405, the available bandwidth of egress tunnels going to the remote domain is determined. At 1410, the egress tunnel having an amount of bandwidth most appropriate for the required level of service is selected.

According to one embodiment of the present invention, an egress tunnel is selected based upon available bandwidth and IP header ToS bits. This is illustrated in FIG. 15. At 1500, the ToS bits for an IP header associated with a session is examined.

At 1505, the available bandwidth of egress tunnels going to the remote domain is determined. At 1510, the egress tunnel having an amount of available bandwidth most appropriate for the required level of service is selected.

Similarly, a VPI/VCI pair may also be used in conjunction with available bandwidth to select an egress tunnel in accordance with another embodiment of the present invention. Referring to FIG. 16, at 1600, a VPI/VCI pair associated with a session is examined. At 1605, a level of service is determined based upon the VPI/VCI pair. At 1610, the available bandwidth of egress tunnels going to the remote domain is determined. At 1615, the egress tunnel having an amount of available most appropriate for the level of service is selected.

According to another embodiment of the present invention, an egress tunnel is selected to provide a relatively even distribution of sessions among egress tunnels, regardless of egress tunnel capacity. This is illustrated in FIG. 17. At 1700, the number of active sessions for each egress tunnel to the remote domain is determined. At 1705, the egress tunnel having the smallest number of active sessions is selected. For example, suppose egress tunnels A and B go to the remote domain and that egress tunnel A has 10 active sessions while egress tunnel B has 5 active sessions. According to this embodiment of the present invention, each new session is added to egress tunnel B until egress tunnel B has more active sessions than egress tunnel A.

According to another embodiment of the present invention, an ingress LAC is instructed to route incomming sessions on a tunnel to an alternate egress LAC when a primary egress LAC exceeds a predetermined loading level. This is illustrated below with reference to FIG. 18.

Turning now to FIG. 18, a flow diagram that illustrates a method for diverting tunneled sessions away from an overloaded egress LAC in accordance with one embodiment of the present invention is presented. At 1800, egress LAC loading is periodically assessed. At 1805, a determination is made regarding whether the egress LAC loading has exceeded a predetermined threshold. If the egress LAC has exceeded the predetermined threshold, at 1810, an ingress LAC is instructed to route incoming sessions on an ingress tunnel to an alternate egress LAC, thus providing relatively efficient utilization of resources and reducing the likelihood of violating a SLA.

A LAC having the capabilities described above can be further used to allow the NSP to behave in a fundamentally different manner than before. Now the NSP can differentiate among computer users. For example, in the case of Joe@ISP.net who is an authorized user of CorpA, the NSP can identify the user at the time the NSP receives his session, look up the details of the SLA with that user, and determine which egress tunnel to the remote domain can best service the session given the requirements of the SLA.

An example of this mode of operation is now presented with reference to FIG. 3. A PoP (Point of Presence) of a NAP 312 is contacted by a user via a PPP connection or

another suitable connection. The PPP session is directed to a NSP 312. The NSP 312 seeks to identify the user with the local AAA server. Instead of performing standard authentication, the AAA server can now perform a query to its tunnel database 386 to determine the "best" egress tunnel that fits the SLA between the NAP 312 and the NSP 310 as well as the SLA between the NSP 310 and the remote domain 300. Assuming that an egress tunnel between the NSP 310 and the remote domain 300 fits the parameters, the session is allocated to the egress tunnel.

Dynamically mapping ingress tunnels to egress tunnels in accordance with embodiments of the present invention also increases scalability by allowing a NSP to share an egress LAC among multiple remote domains without manually reconfiguring the network.

While embodiments and applications of this invention have been shown and described, it would be apparent to those skilled in the art having the benefit of this disclosure that many more modifications than mentioned above are possible without departing from the inventive concepts herein. The invention, therefore, is not to be restricted except in the spirit of the appended claims.